# Lecture 6: PCA in high dimensions, random matrix theory and financial applications

## Foundations of Data Science: Algorithms and Mathematical Foundations

Mihai Cucuringu
mihai.cucuringu@stats.ox.ac.uk

CDT in Mathematics of Random System
University of Oxford

21 September, 2023

Slides by Jim Gatheral, *Random Matrix Theory and Covariance Estimation* (with minor changes) + *Financial Applications of Random Matrix Theory: a short review*, J.P. Bouchaud and M. Potters

# Motivation

- some of the state-of-the-art optimal liquidation portfolio algos (balancing risk vs impact cost) involve inverting the covariance matrix
- eigenvalues of the covariance matrix that are small (or even zero) correspond to portfolios of stocks that have
  - nonzero returns
  - extremely low or vanishing risk
- such portfolios are invariably related to estimation errors resulting from insuffient data
- use random matrix theory to alleviate the problem of small eigenvalues in the estimated covariance matrix

Goal is to understand:

- the basis of random matrix theory (RMT)
- how to apply RMT to estimating covariance matrices
- performance comparison with baselines

# Roadmap

- A financial interpretation
- Random matrix theory
  - Random matrix examples
  - Wigner's semicircle law
  - The Marčenko-Pastur density
  - The Tracy-Widom law
  - Impact of fat tails
- Estimating correlations
  - Uncertainty in correlation estimates
  - Example with SPX stocks
  - Recipe for filtering the sample correlation matrix
- *Comparison with Barra*
  - *Comparison of eigenvectors*
  - *The minimum variance portfolio*
    - *Comparison of weights*
    - *In-sample and out-of-sample performance*
- Approaches to covariance matrix estimation

# A financial interpretation

- let $N$ denote number of stocks
- let $T$ denote the number of observations (eg, daily returns)
- let $E$ denote the Pearson estimator of the correlation matrix

$$E_{ij} = \frac{1}{T}\sum_{t=1}^{T} r_i^t r_j^t \equiv (X^T X)_{ij} \tag{1}$$

(the empirical correlation matrix, on a given realization)

- $X_{ti} = \frac{r_i^t}{\sqrt{T}}$, where $r_i^t$ is the realization of quantity $i \in \{1, \ldots, N\}$ at time $t \in \{1, \ldots, T\}$ (already demeaned and standardized).
- for fixed $N$ and $T \to \infty$, all eigenvalues and their corresponding eigenvectors can be "trusted" to extract meaningful information
- however, not the case if $q = \frac{N}{T} = O(1)$, when only a subset of the eigen-spectrum of the *true* corr. mtx. $C$ can be reliably estimated
- since $E$ is by construction rank deficient, $(N - T)$ eigenvalues are exactly equal to zero (clearly these spurious eigenvalues do not correspond to any real structure in $C$)
- useful to give a physical/financial interpretation of the eigenvectors $\vec{V}^{(k)}$, $k = 1, 2, \ldots$

# A financial interpretation

- consider the eigenvectors $\vec{V}^{(k)}$, $k = 1, 2, \ldots$ of $E$
- interpret the entries of the eigenvector as $\vec{V}_1^{(k)}, \vec{V}_2^{(k)}, \ldots, \vec{V}_N^{(k)}$ as the weights of the different stocks $i = 1, \ldots, N$ in a certain portfolio $\Pi_k$, where
  - some stocks are *"long"* (i.e, $(V_i^{(k)} > 0)$
  - while others are *"short"* (i.e, $(V_i^{(k)} < 0)$
- the realized risk $\mathcal{R}_k^2$ of portfolio $\Pi_k$, as measured by the variance of its returns, is given by

$$\mathcal{R}_k^2 = \frac{1}{T} \sum_{t=1}^{T} \left( \sum_{i=1}^{N} \vec{V}_i^{(k)} r_i^t \right)^2 = \sum_{ij} \vec{V}_i^{(k)} \vec{V}_j^{(k)} E_{ij} =: \lambda_k \qquad (2)$$

- note the last term is simply the quadratic form $x^T E x$ (denoting any eigenvector by $x$), hence equal to an eigenvalue
- also note $\vec{V}_i^{(k)} r_i^t$ is essentially the PnL (Profit and Loss) obtained from investing $\vec{V}_i^{(k)}$ notional amount ($$$) in a stock whose return on day $t$ is $r_i^t$
  - if $\vec{V}_i^{(k)}$ and $r_i^t$ have the same sign (both +ve or both -ve), you win
  - otherwise it's a losing bet

# A financial interpretation

- eigenvalue $\lambda_k$ gives the risk of investing in portfolio $\Pi_k$
- large eigenvalues correspond to a risky mix of assets
- small eigenvalues correspond to a particularly quiet/less volatile mix of assets
- typically, when looking at equity data (stock markets), the largest eigenvalue corresponds to investing roughly equally on all stocks $V_i^1 \approx \frac{1}{\sqrt{N}}$ (or perhaps proportional to market cap)
- called the ***market mode*** - strongly correlated with the market index (eg, $S\&P$ 500)
- there is no diversification in this portfolio: the only bet is whether the market as a whole will go up or down (difficult task btw), hence the risk being large
- conversely, if two stocks move very tightly together (canonical example: Coca-cola & Pepsi): buying one and selling the other
  - leads to a portfolio that barely moves
  - sensitive only to events that strongly differentiate the 2 companies
  - there corresponds a small eigenvalue of $E$ with an eigenvector that is very localized, eg, $(0, 0, \ldots, 0, \sqrt{2}/2, 0, \ldots, -\sqrt{2}/2, 0, \ldots, 0, 0)$

# Uncorrelated eigenportfolios

- another property of the eigenvector portfolios $\Pi_k$ is that their returns are uncorrelated

$$\frac{1}{T}\sum_{t=1}^{T}\left(\sum_{i=1}^{N}\vec{V}_i^{(k)}r_i^t\right)\left(\sum_{j=1}^{N}\vec{V}_i^{(l)}r_j^t\right)=\sum_{ij}\vec{V}_i^{(k)}\vec{V}_j^{(l)}E_{ij}=\lambda_k\delta_{k,l} \quad (3)$$

  where $\delta_{k,l}$ denote the Kronecker-delta

- performing PCA on the empirical correlation matrix $E$ provides a list of ***eigen-portfolios***, $\Pi_1, \Pi_2, \ldots, \Pi_k$, corresponding to uncorrelated investments, sorted in decreasing variance

- Q: when traversing the spectrum, how far down should we go, before everything becomes just noise?

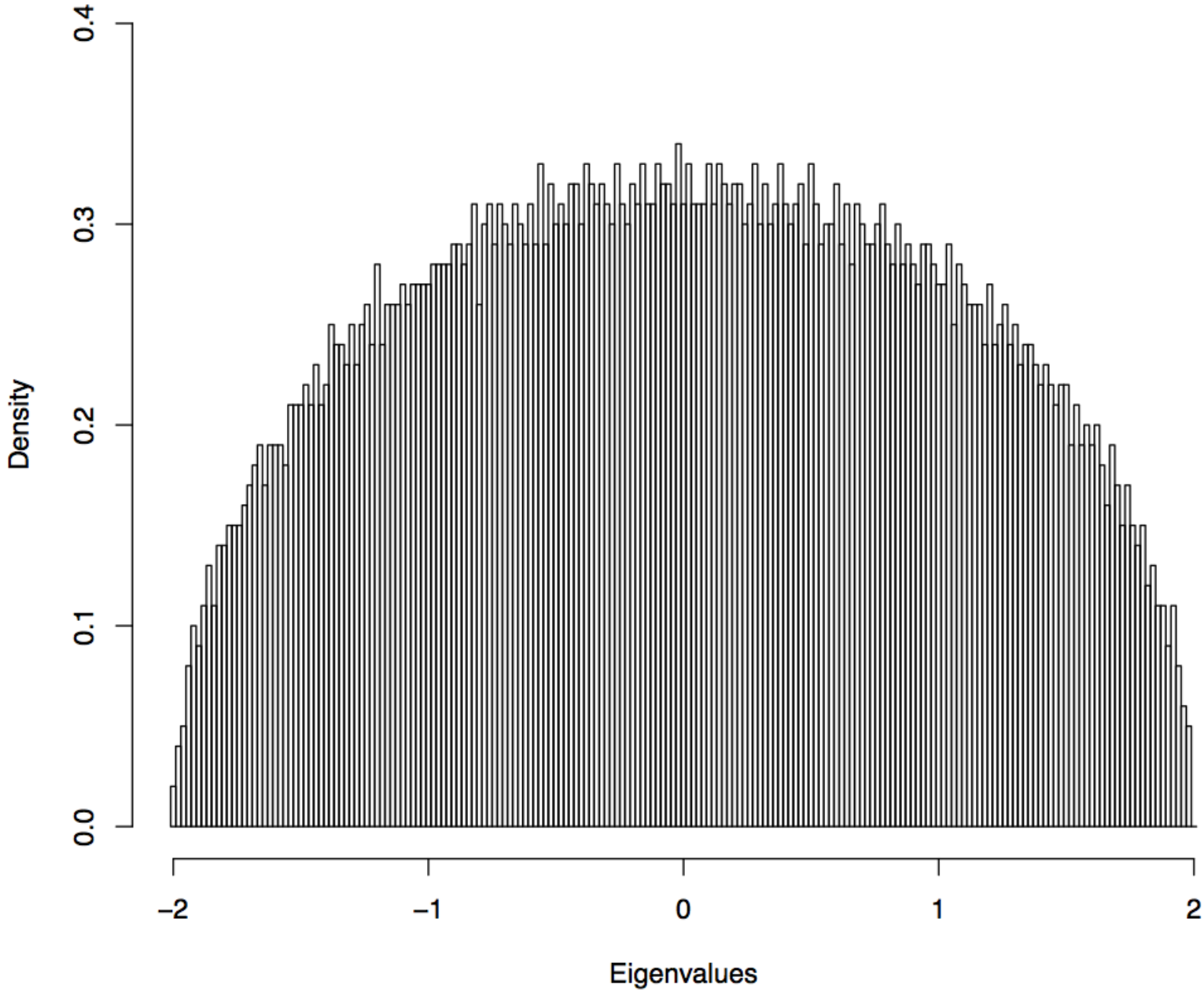
# Example 1: Normal random symmetric matrix

- Generate a $5,000 \times 5,000$ random symmetric matrix with entries $A_{ij} \sim N(0,1)$
- Compute eigenvalues
- Draw the histogram of all eigenvalues.

R-code to generate a symmetric random matrix whose off-diagonal elements have variance $\frac{1}{N}$:

- n = 5000
- m = array(rnorm($n^2$),c(n,n));
- m2 = (m+t(m))/sqrt(2*n); # Make m symmetric
- lambda = eigen(m2, symmetric=T, only.values = T);
- ev = lambda$values;
- hist(ev, breaks=seq(-2.01,2.01,0.02),main=NA, xlab="Eigenvalues",freq=F)

# Normal random symmetric matrix



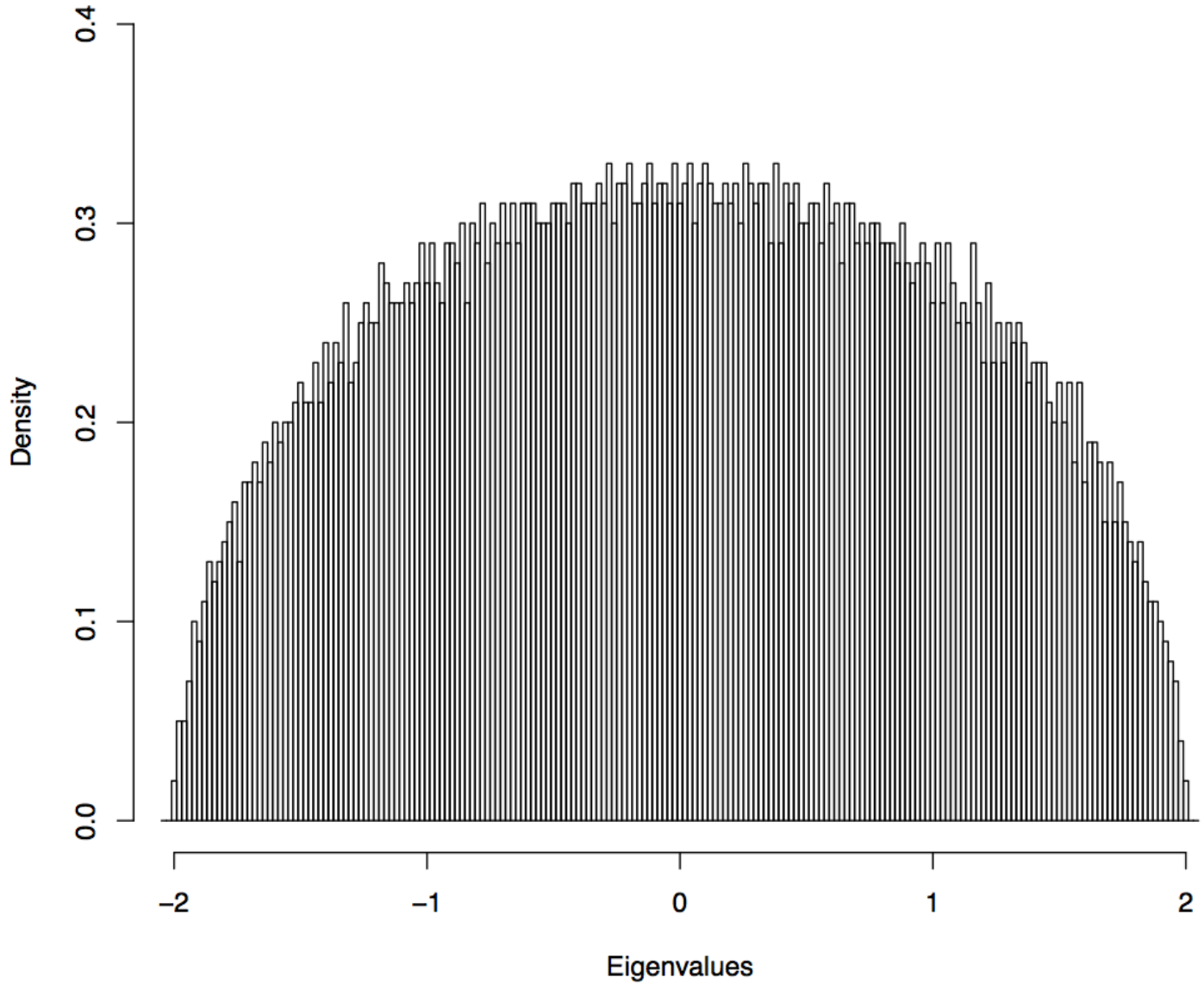
# Uniform random symmetric matrix

- Generate a 5,000 x 5,000 random symmetric matrix with entries $A_{ij} \sim \text{Uniform}(0, 1)$
- Compute eigenvalues
- Draw the histogram of all eigenvalues

R-code:

- n = 5000;
- mu = array(runif($n^2$),c(n,n))
- mu2 = sqrt(12)*(mu+t(mu)-1)/sqrt(2*n)
- lambdau = eigen(mu2, symmetric=T, only.values = T)
- ev = lambdau$values;
- hist(ev, breaks=seq(-2.01,2.01,0.02), main=NA, xlab="Eigenvalues",freq=F)

# Uniform random symmetric matrix




Striking pattern: the density of eigenvalues is still a semicircle!

# Wigner's semicircle law

Let $\tilde{A}$ by an $N \times N$ matrix with entries $\tilde{A}_{ij} \sim N(0, \sigma^2)$. Define

$$A_N = \frac{1}{\sqrt{2N}} \left( A + A^T \right)$$

- $A_N$ is symmetric with variance

$$\text{Var}[a_{ij}] = \begin{cases} \sigma^2/N & \text{if } i \neq j \\ 2\sigma^2/N & \text{if } i = j \end{cases} \tag{4}$$

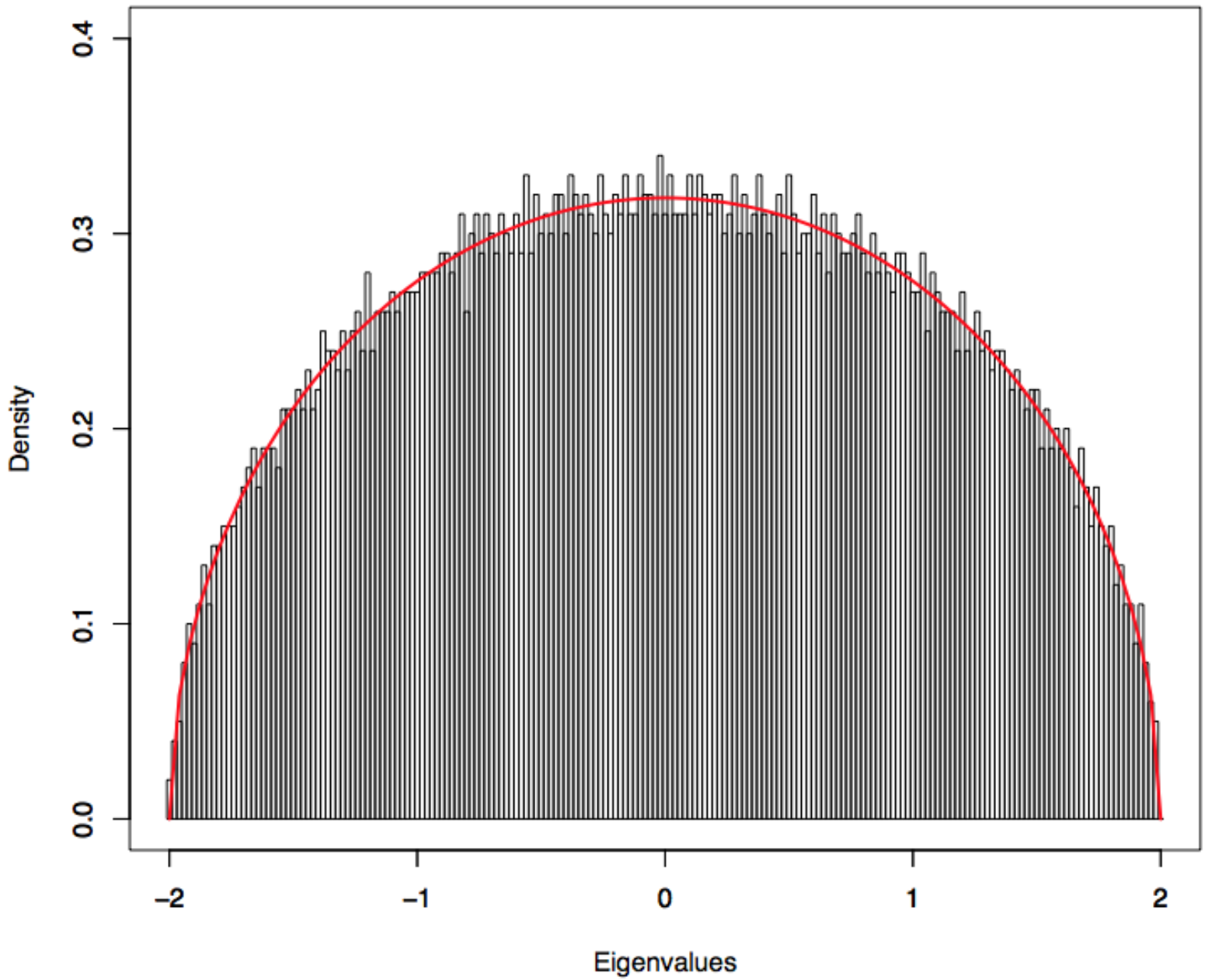
- the density of eigenvalues of $A_N$ is given by

$$\rho_N(\lambda) \stackrel{\text{def}}{=} \frac{1}{N} \sum_{i=1}^{N} \delta(\lambda - \lambda_i)$$

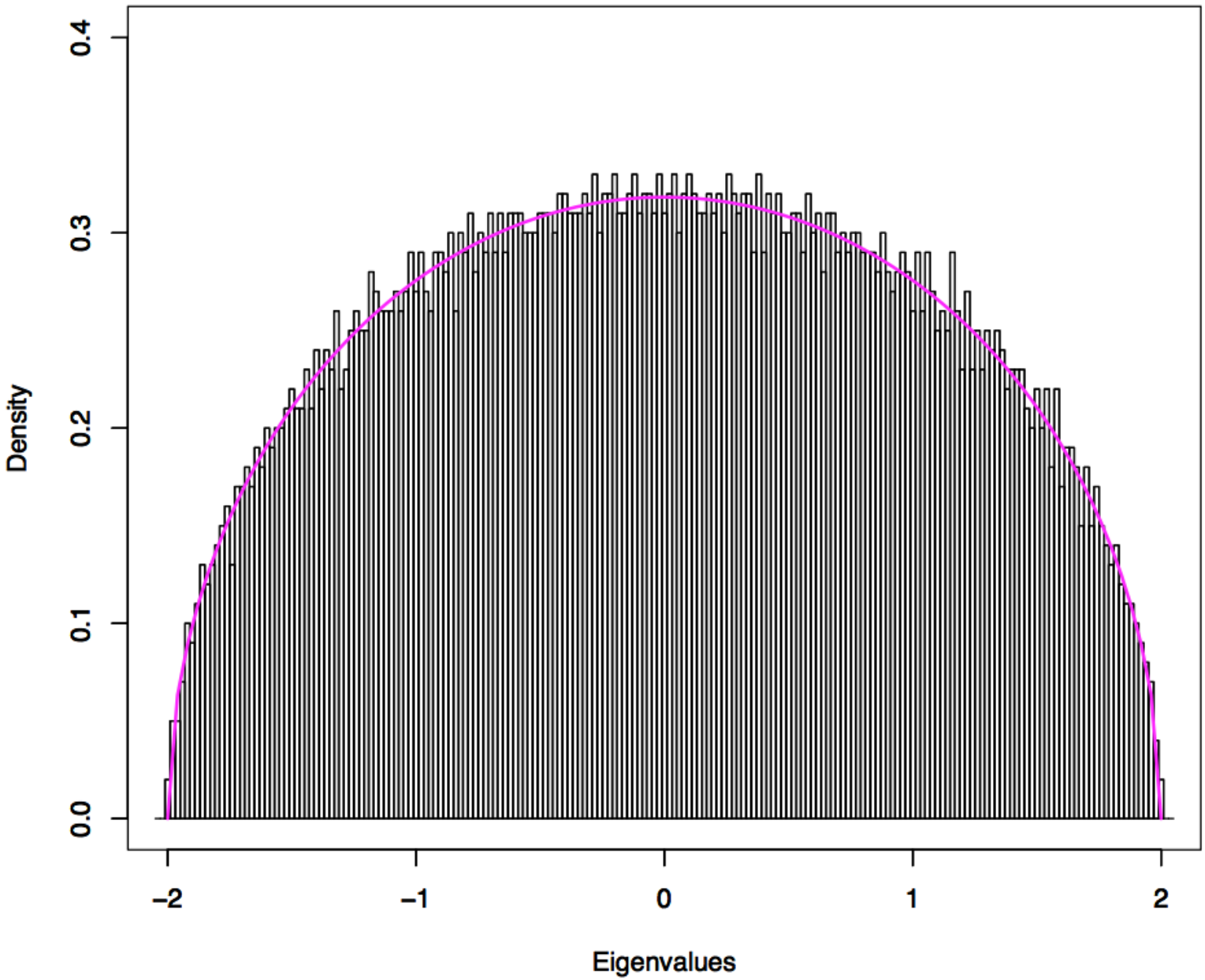
(empirical spectral distribution), which, as shown by Wigner

$$\text{as } n \to \infty, \rho_N(\lambda) \longrightarrow \begin{cases} \frac{1}{2\pi\sigma^2} \sqrt{4\sigma^2 - \alpha^2} & \text{if } |\lambda| \leq 2\sigma \\ 0 & \text{otherwise} \end{cases} \stackrel{\text{def}}{=} \rho(\lambda) \tag{5}$$

# Normal random matrix + Wigner semicircle density



# Uniform random matrix + Wigner semicircle density



# Random correlation matrices

- we have $M$ stock return series with $T$ elements each
- the elements of the $M \times M$ empirical correlation matrix E are

$$E_{ij} = \frac{1}{T} \sum_{t=1}^{T} x_{it} x_{jt}$$

where $x_{it}$ denotes the return at time $t$ of stock $i$, normalized by the standard deviation so that $\mathrm{Var}[x_{it}] = 1$

- in compact matrix form, this can be written as

$$E = HH^T$$

where $H$ is the $M \times T$ matrix whose rows are the time series of returns, one for each stock (demeaned and standardized)

# Eigenvalue spectrum of random correlation matrix

- Suppose the entries of $H$ are random with variance $\sigma^2$
- Then, in the limit T, M $\to \infty$, while keeping the ratio $Q \stackrel{\text{def}}{=} \frac{T}{M} \geq 1$ constant, the density of eigenvalues of $E$ is given by

$$\rho(\lambda) = \frac{Q}{2\pi\sigma^2} \frac{\sqrt{(\lambda_+ - \lambda)(\lambda_- - \lambda)}}{\lambda}$$

where the max and min eigenvalues are given by

$$\lambda_\pm = \sigma^2 \left(1 \pm \sqrt{\frac{1}{Q}}\right)^2$$

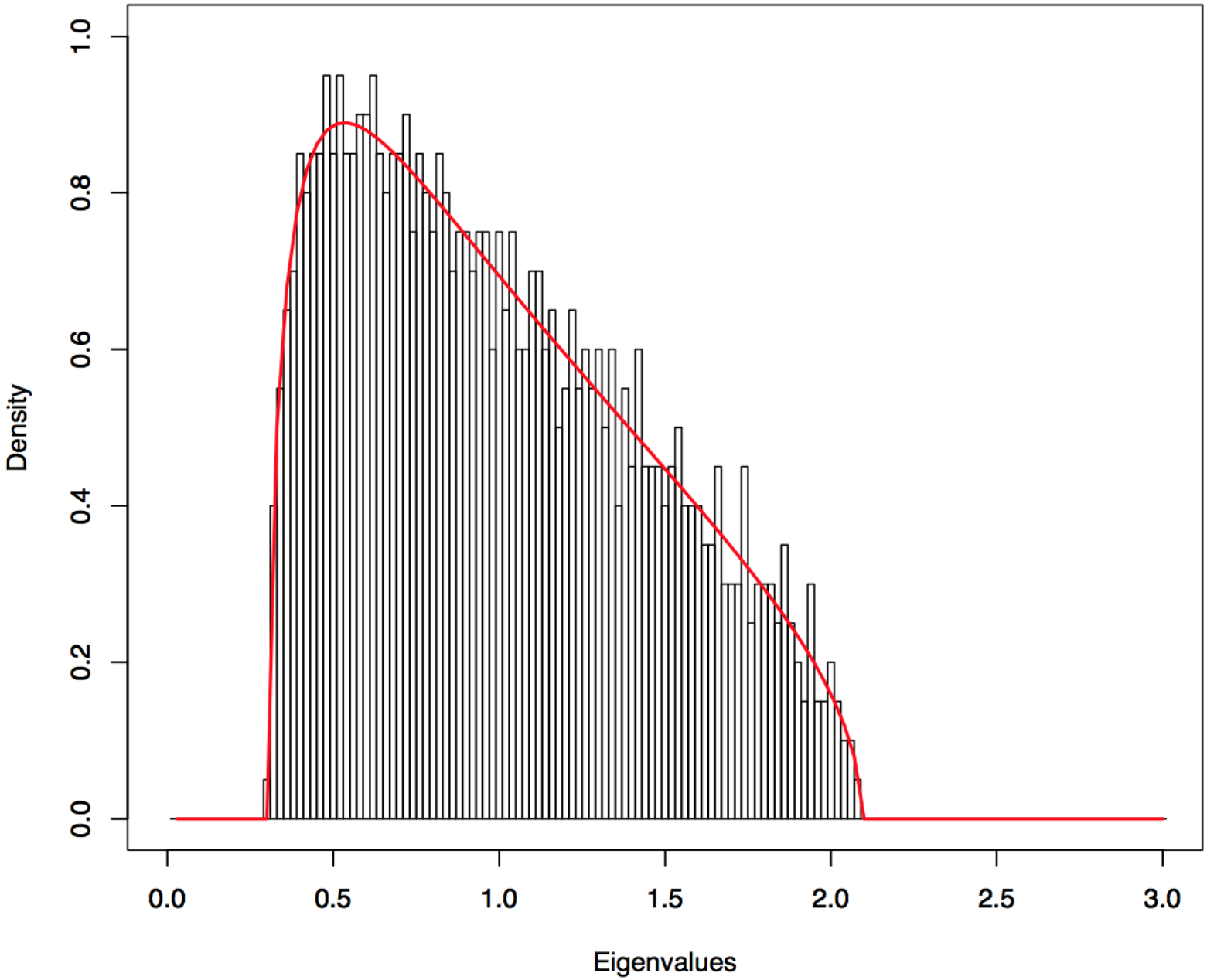
- $\rho(\lambda)$ is also known as the Marčhenko-Pastur distribution that describes the asymptotic behavior of eigenvalues of large random matrices

# Example: IID random normal returns

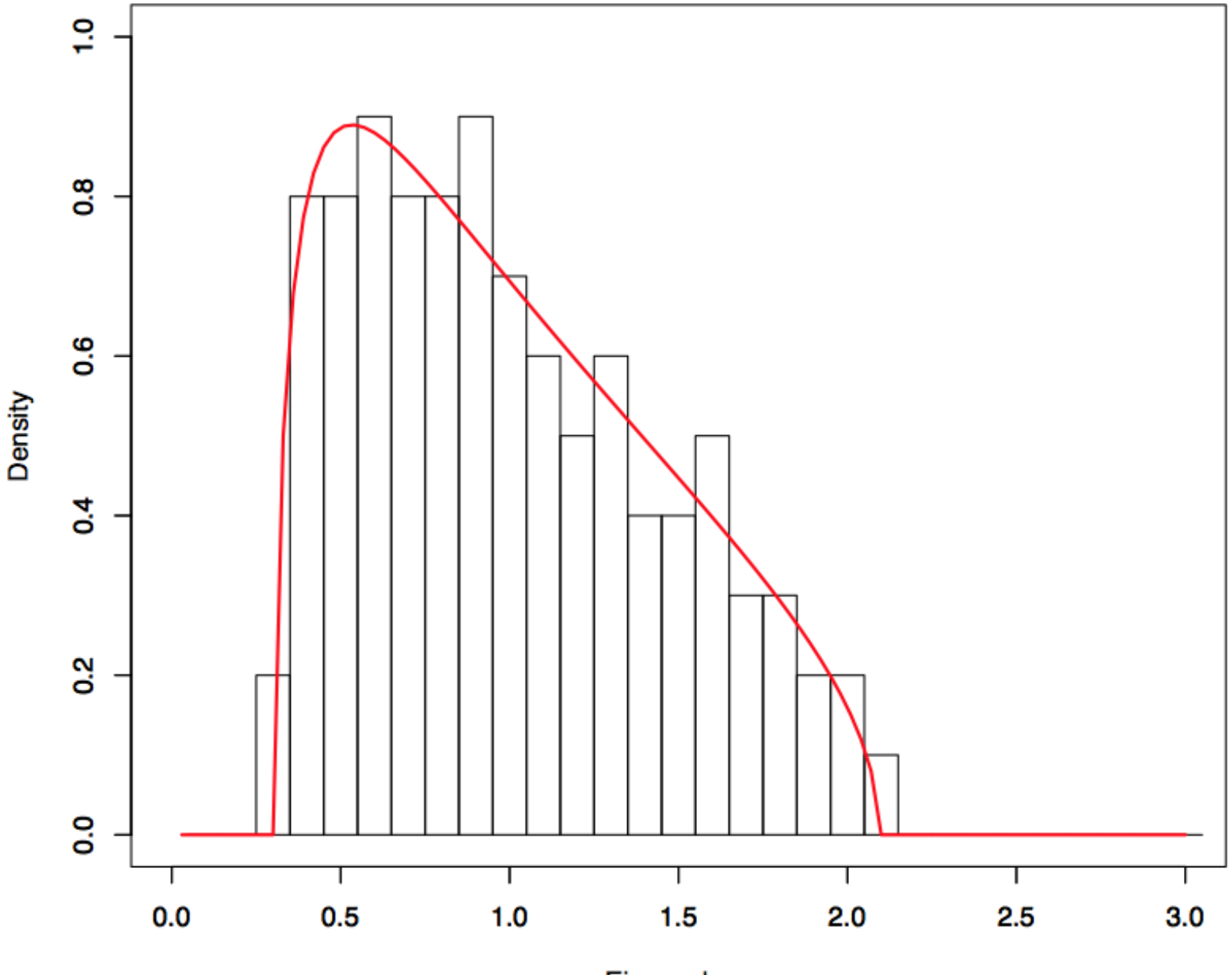
R-code:

- t = 5000;
- m = 1000;
- h = array(rnorm(m*t),c(m,t)); # Time series in rows
- e = h % * % t(h)/t; # Form the correlation matrix
- lambdae = eigen(e, symmetric=T, only.values = T);
- ee = lambdae$values;
- hist(ee, breaks =seq(0.01,3.01,.02), main=NA, xlab=”Eigenvalues”, freq=F)

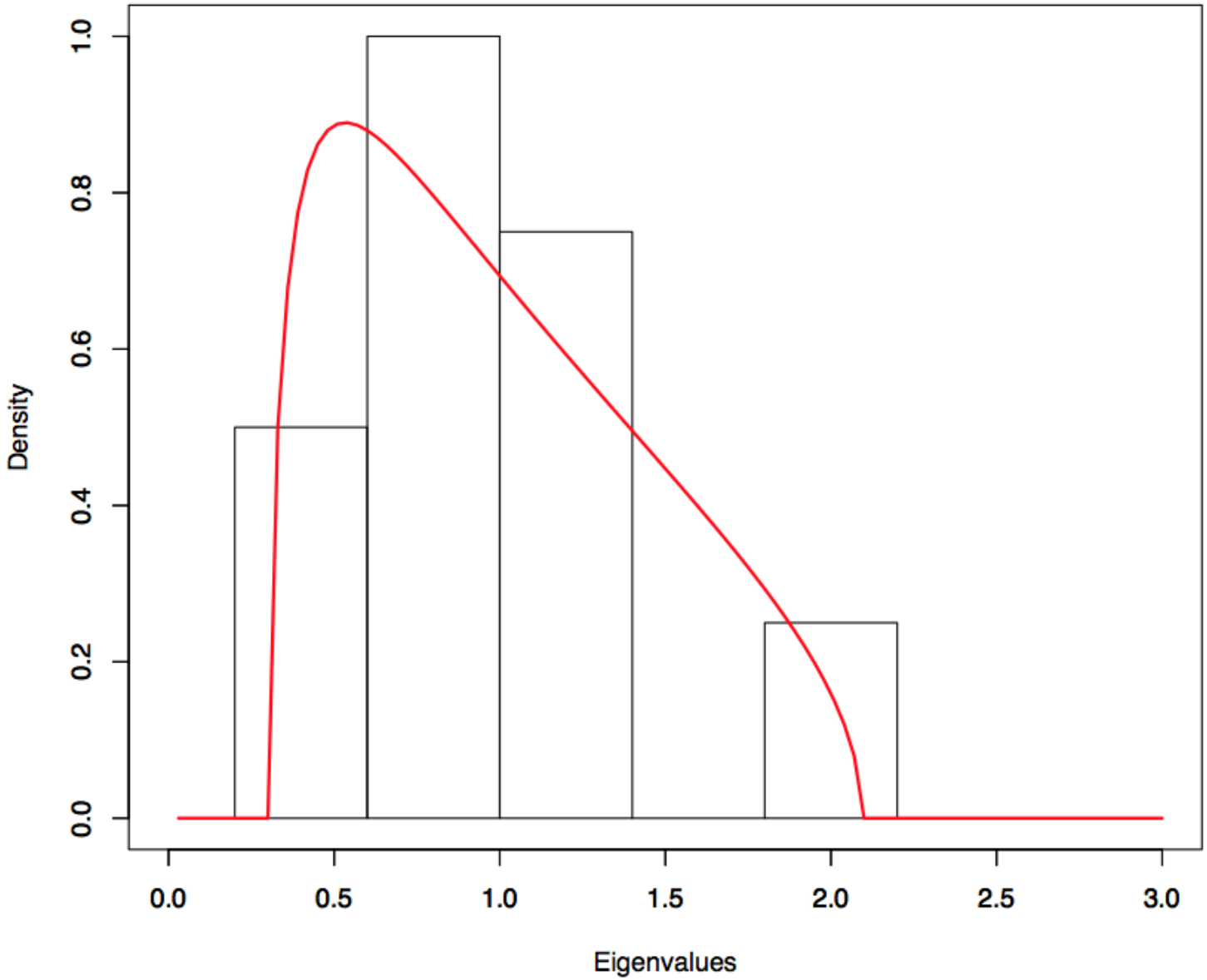
# Empirical density with superimposed Marčhenko-Pastur density



# Empirical density for $M = 100$, $T = 500$ (with Marčenko-Pastur density superimposed)

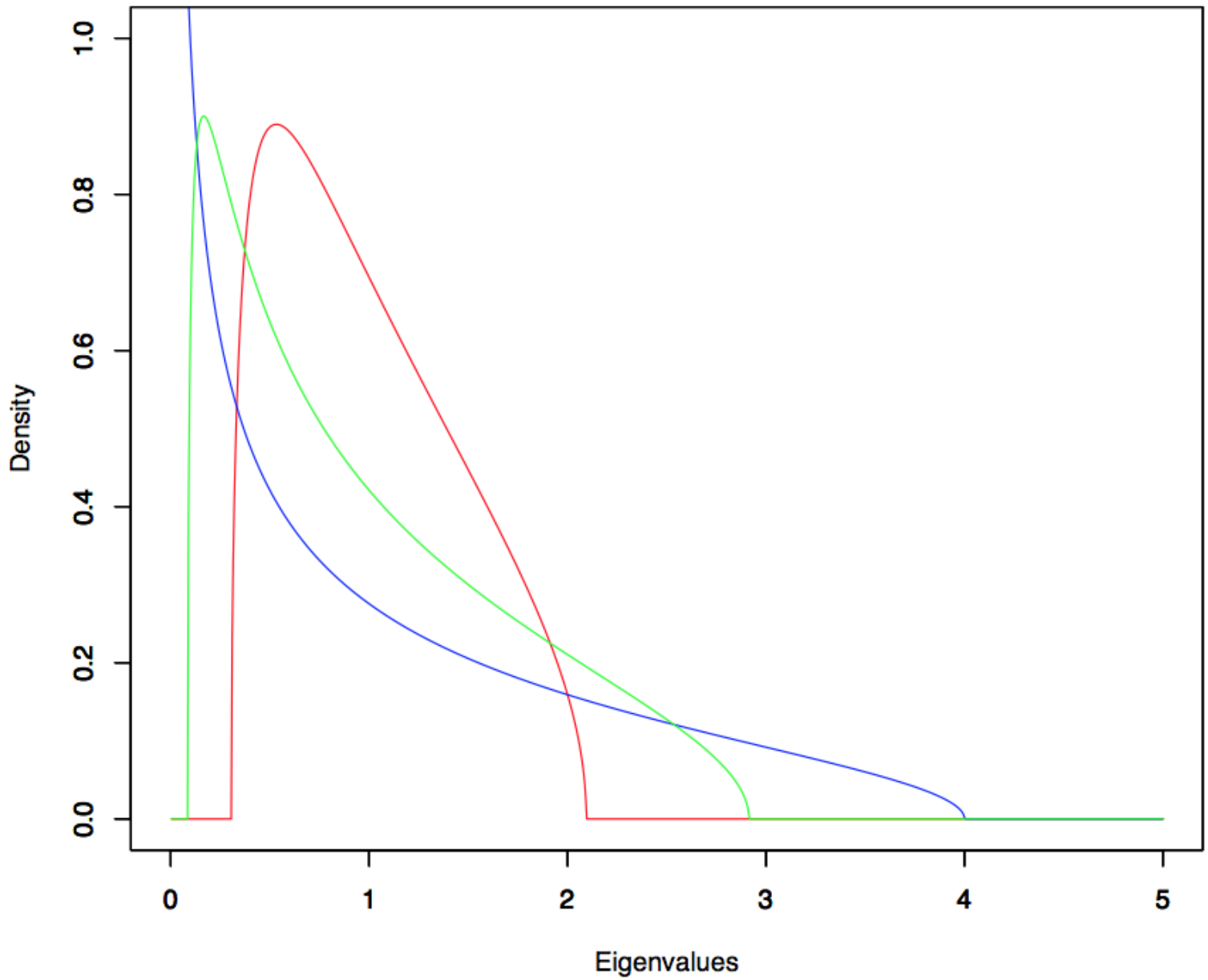


# Empirical density for $M = 10$, $T = 50$ (with Marčenko-Pastur density superimposed)



# Marčenko-Pastur densities depends on $Q = T/M$

Density for $Q = \{$ 1 (blue), 2 (green), 5 (red) $\}$



# Tracy-Widom: of the largest eigenvalue

- For certain applications, we would like to know
  - where the random bulk of eigenvalues ends
  - where the spectrum of eigenvalues corresponding to true information begins
- $\Rightarrow$ need to know the distribution of the largest eigenvalue
- The distribution of the largest eigenvalue of a random correlation matrix is given by the Tracy-Widom law

$$P(T\lambda_{max} < \mu_{TM} + s\sigma_{TM}) = F_1(s)$$

where

$$\mu_{TM} = \left(\sqrt{T - \frac{1}{2}} + \sqrt{M - \frac{1}{2}}\right)^2$$

$$\sigma_{TM} = \left(\sqrt{T - \frac{1}{2}} + \sqrt{M - \frac{1}{2}}\right)\left(\frac{1}{\sqrt{T - \frac{1}{2}}} + \frac{1}{\sqrt{M - \frac{1}{2}}}\right)^{1/3}$$

On the top eigenvalue of heavy-tailed random matrices G. Biroli, J.-P. Bouchaud, and M. Potters
https://iopscience.iop.org/article/10.1209/0295-5075/78/10001/fulltext/

# Fat-tailed random matrices

So far, we have considered matrices whose entries are

- Gaussian
- uniformly distributed

But, in practice: stock returns exhibit a fat-tailed distribution

- Bouchaud et al.: fat tails can massively increase the maximum eigenvalue in the theoretical limiting spectrum of the random matrix
- ”Financial Applications of Random Matrix Theory: a short review”, J.P. Bouchaud and M. Potters
  `http://arxiv.org/abs/0910.1205`
- For extremely fat-tailed distributions (Cauchy for example), the semi-circle law no longer holds

# Sampling error

- suppose we compute the sample correlation matrix of M stocks with T returns in each time series.
- assume the true correlation were the identity matrix
- Q: expected value of the greatest sample correlation entry?
- for $N(0,1)$ distributed returns, the median maximum correlation $\rho_{max}$ should satisfy:

$$\log 2 \approx \frac{M(M-1)}{2} N(-\rho_{max}\sqrt{T})$$

- with M = 500,T = 1000, we obtain $\rho_{max} \approx 0.14$
- sampling error induces spurious (and potentially significant) correlations between stocks!

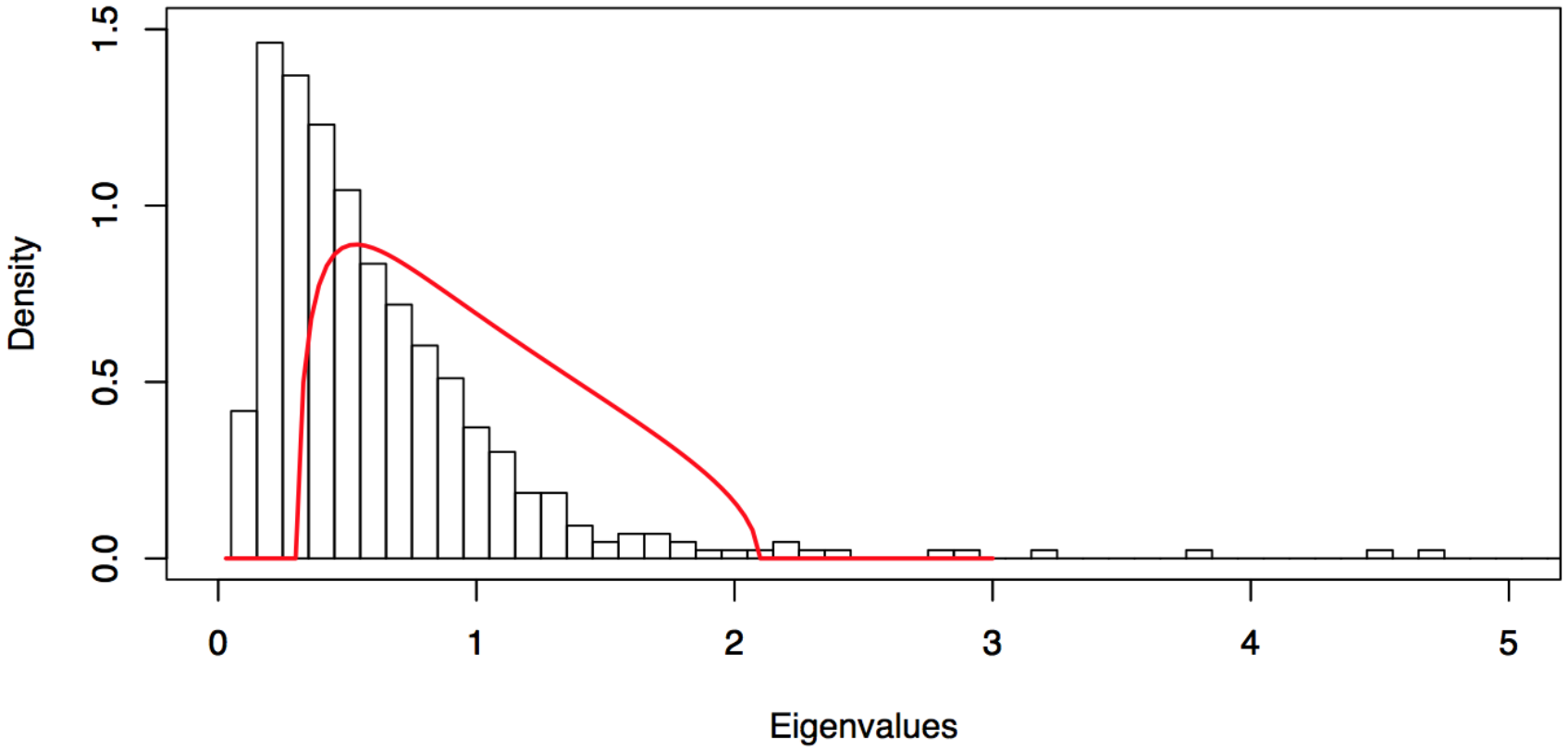
# An experiment with real data

- $M = 431$ stocks in the S&P 500 index for which we have $T = 5 \times 431 = 2155$ consecutive daily returns

- $Q = T/M = 5$
- There are $M(M-1)/2 = 92,665$ distinct entries in the correlation matrix to be estimated from $2,155 \times 431 = 928,805$ data points

First, compute the eigenvalue spectrum and superimpose the Marčenko-Pastur density with $Q = 5$.

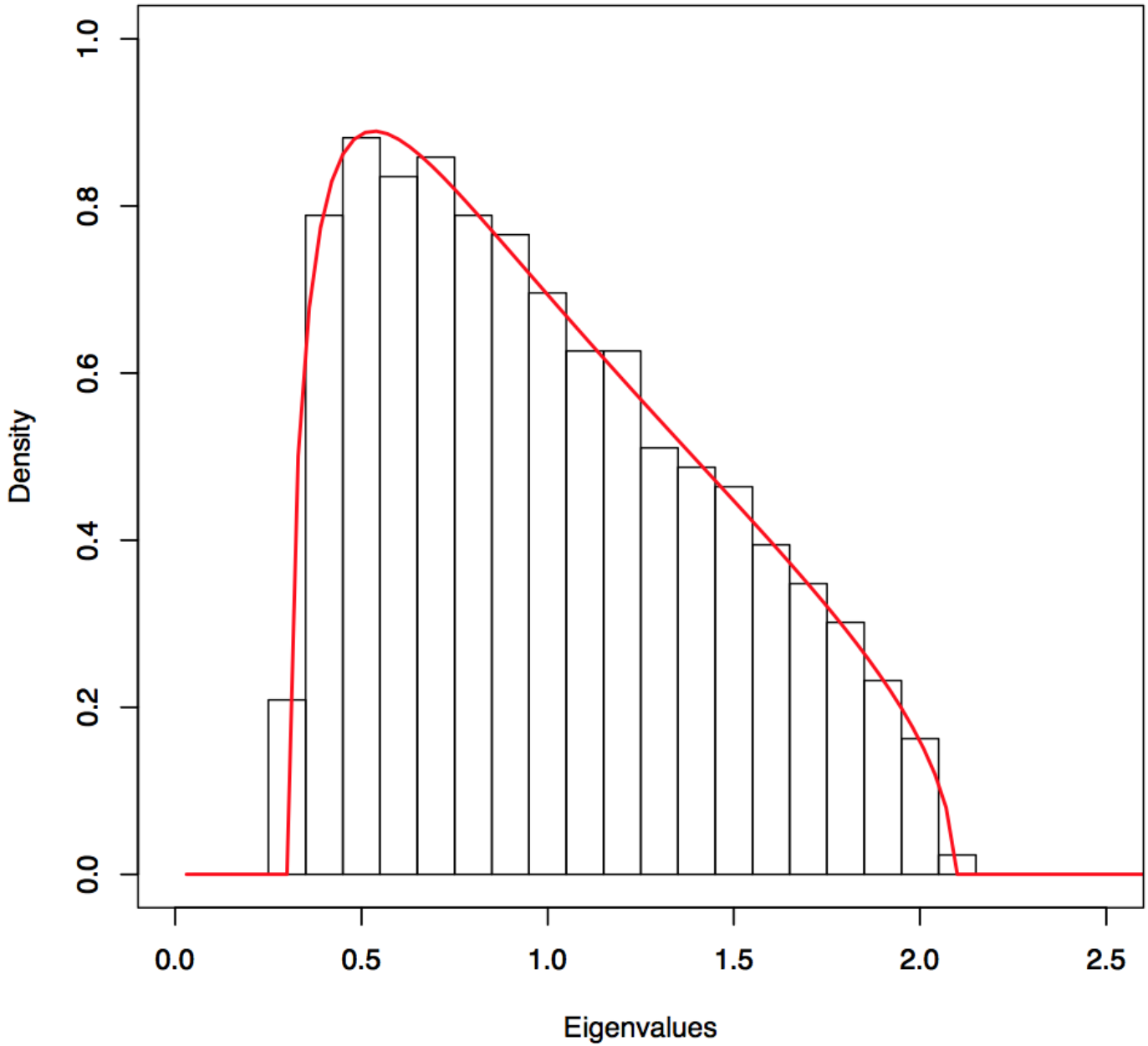
# The eigenvalue spectrum of the sample correlation



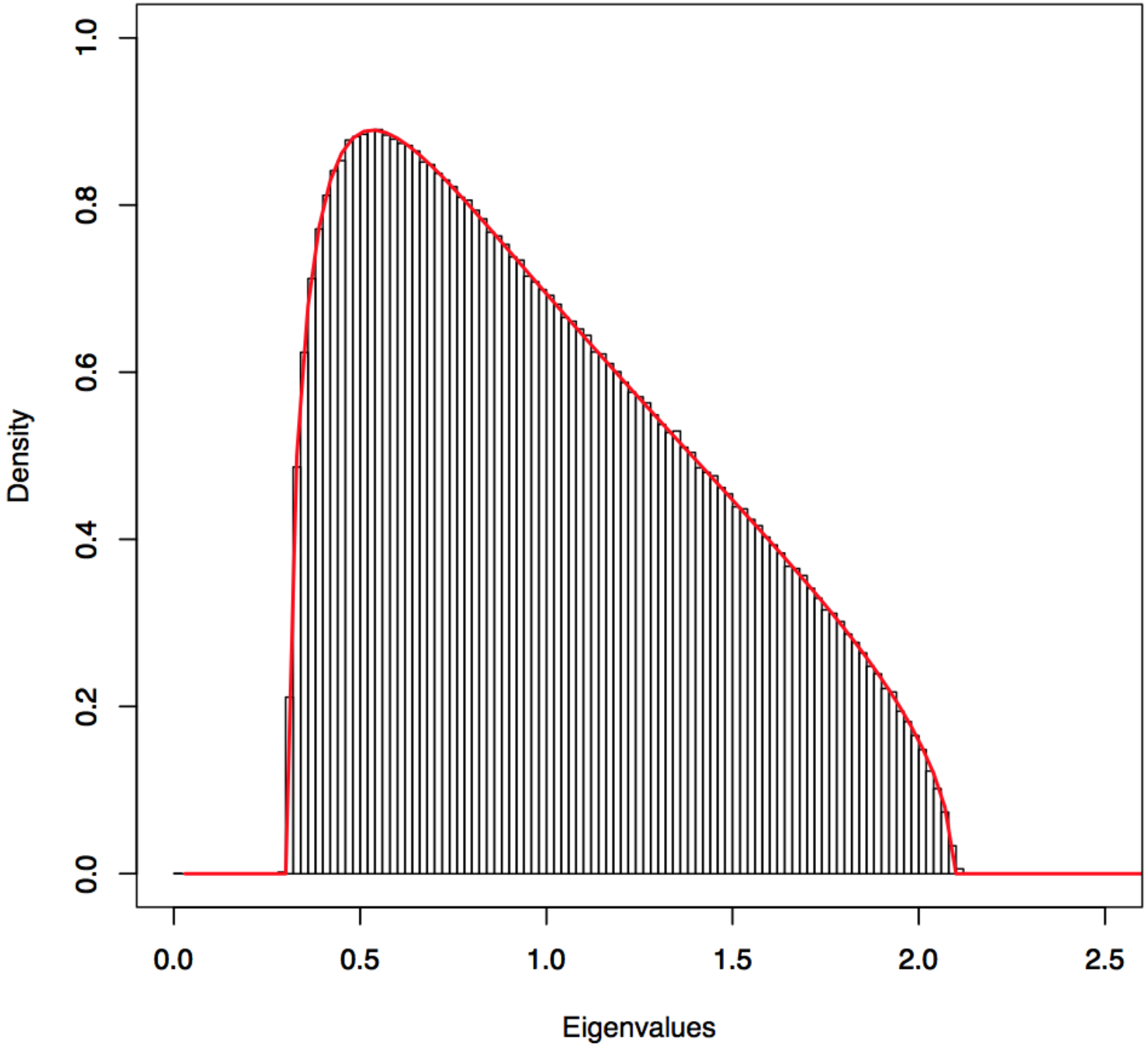

Note that the top eigenvalue is 105.37 (the market) – way off the end of the chart! The next biggest eigenvalue is 18.73. Both were left out for ease of visualization.

# With randomized return data

If we shuffle the returns in each time series:

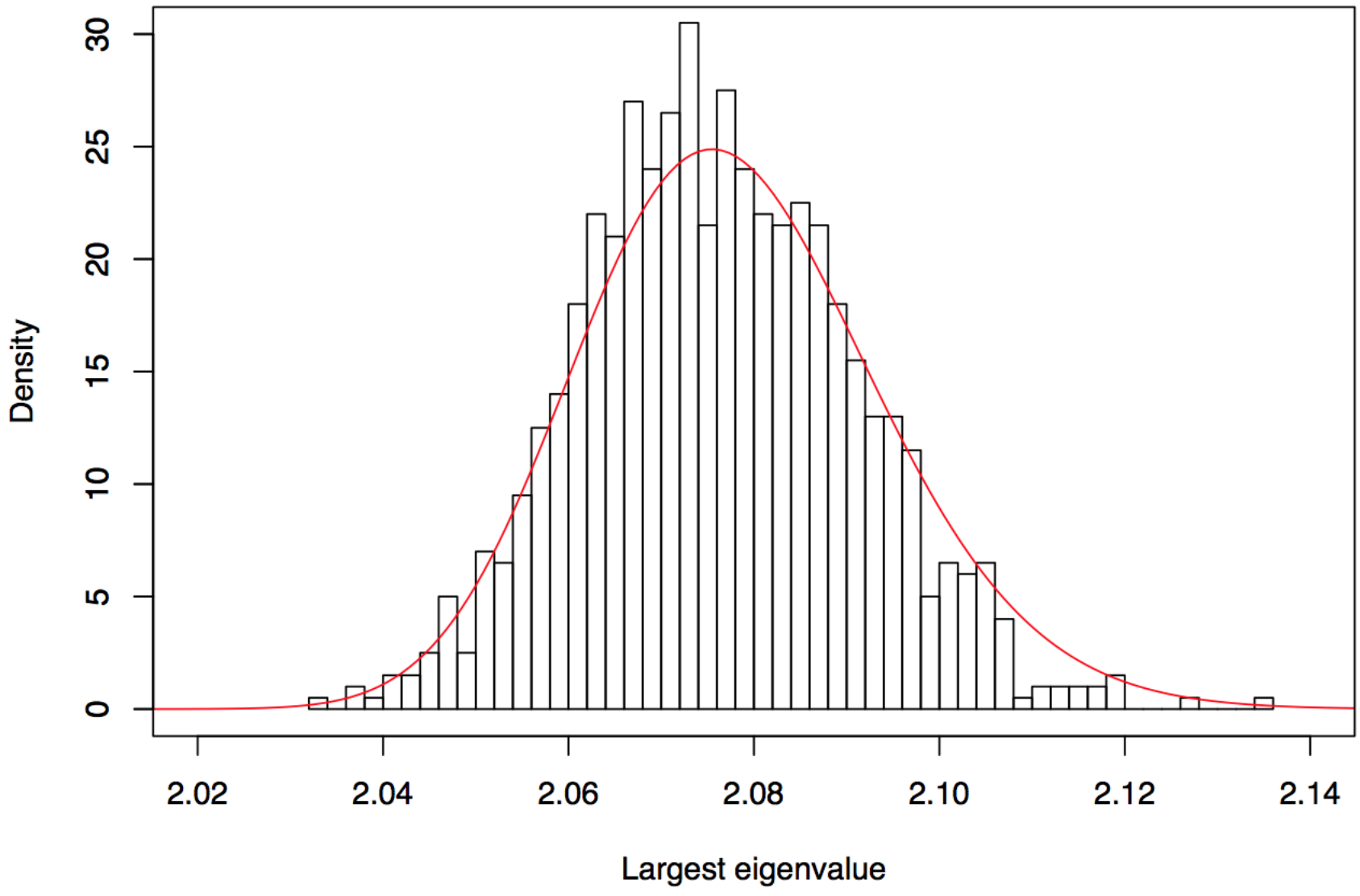


# Repeating 1,000 times and average



# Distribution of the largest eigenvalue

We can compare the empirical distribution of the largest eigenvalue with the Tracy-Widom density (in red):



# Interim conclusions

Remarks:

- Even though return series are fat-tailed:
  - the Marčenko-Pastur density is a very good approximation to the density of eigenvalues of the correlation matrix of the randomized returns
  - the Tracy-Widom density is a good approximation to the density of the largest eigenvalue of the correlation matrix of the randomized returns
- the Marčenko-Pastur density **does not** remotely fit the eigenvalue spectrum of the sample correlation matrix
  - $\Rightarrow$ there is non-random structure in the return data
- can compute the theoretical spectrum arbitrarily accurately by performing numerical simulations

# Problem formulation

- Which eigenvalues are significant and how do we interpret their corresponding eigenvectors?

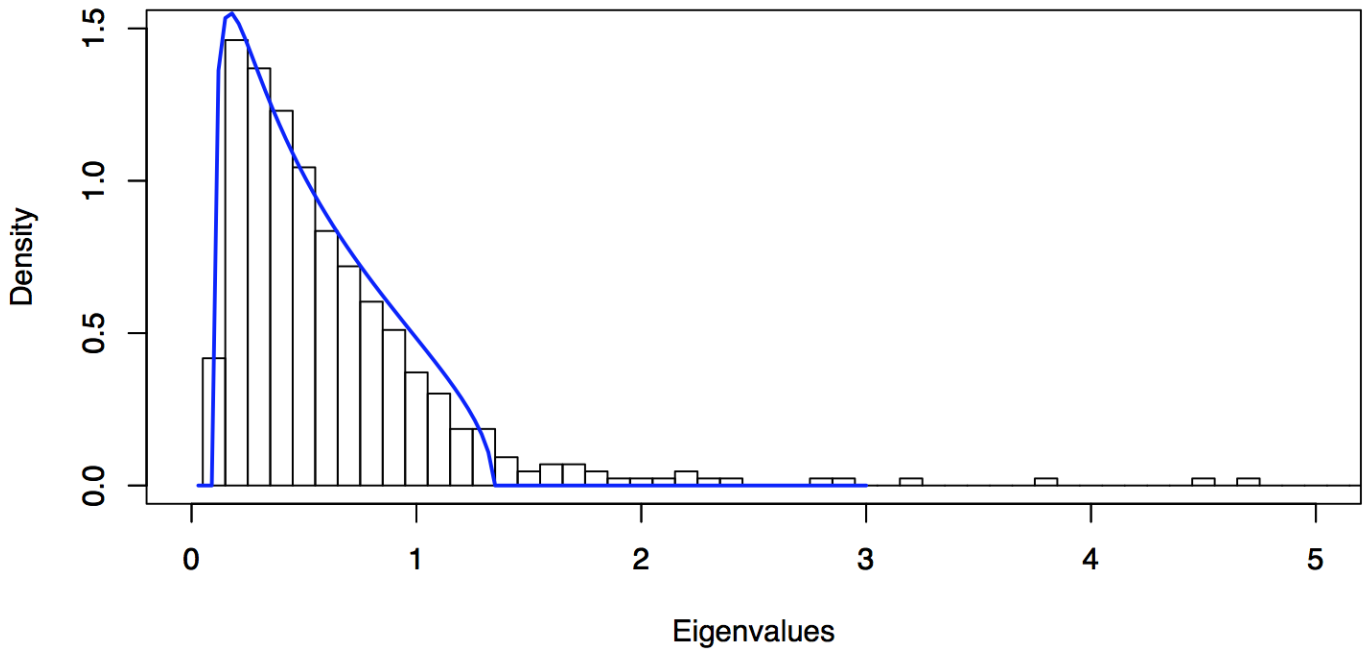
# A hand-waving practical approach

Suppose we find the values of $\sigma$ and $Q$ that best fit the bulk of the eigenvalue spectrum.

We find

$$\sigma = 0.73; Q = 2.9$$

and obtain the following plot:



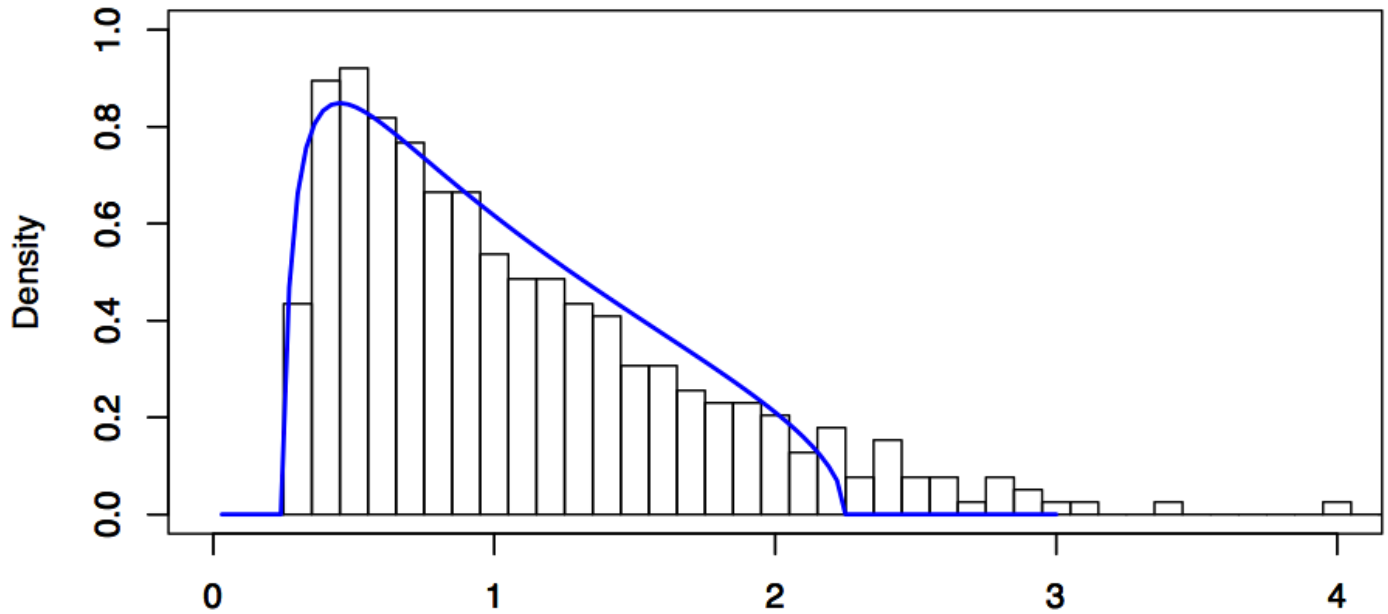

Max and min MP eigenvalues are 1.34 and 0.09 respectively.

# Some empirical analysis

- If we are to believe this estimate, a fraction $\sigma^2 = 0.53$ of the variance is explained by eigenvalues that correspond to random noise. The remaining fraction 0.47 has information
- From the plot, it looks as if we should cut off eigenvalues above 1.5 or so
- Summing the eigenvalues themselves, we find that 0.49 of the variance is explained by eigenvalues greater than 1.5

# More carefully: correlation matrix of residual returns

- For each stock, subtract factor returns associated with the top 25 eigenvalues ($\lambda > 1.6$)
- For $\sigma = 1; Q = 4$ we get the best fit of the Marčenko-Pastur density and obtain the following plot:




- Maximum and minimum Marčenko-Pastur eigenvalues are 2.25 and 0.25 respectively.

# Distribution of eigenvector components

- if there is no information in an eigenvector, we expect the distribution of the components to be a maximum entropy distribution. Which one?
- if we normalized the eigenvector $u$ such that its components $u_i$ satisfy

$$\sum_{i=1}^{M} u_i^2 = M,$$

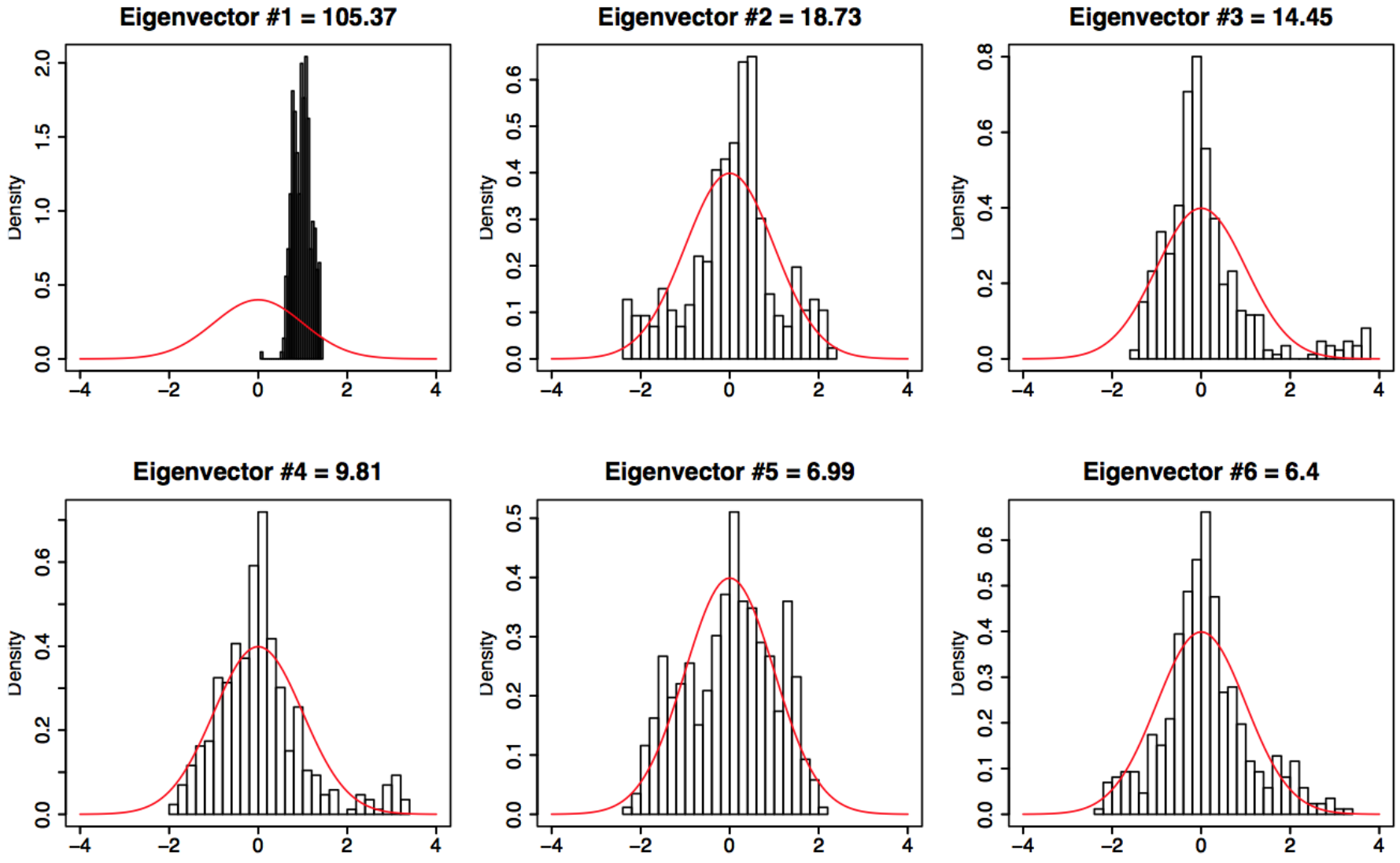
the distribution of the $u_i$ should have the limiting density

$$p(u) = \sqrt{\frac{1}{2\pi}} e^{-\frac{u^2}{2}}$$

- next, superimpose the empirical distribution of eigenvector components and the zero-information limiting density for various eigenvalues...

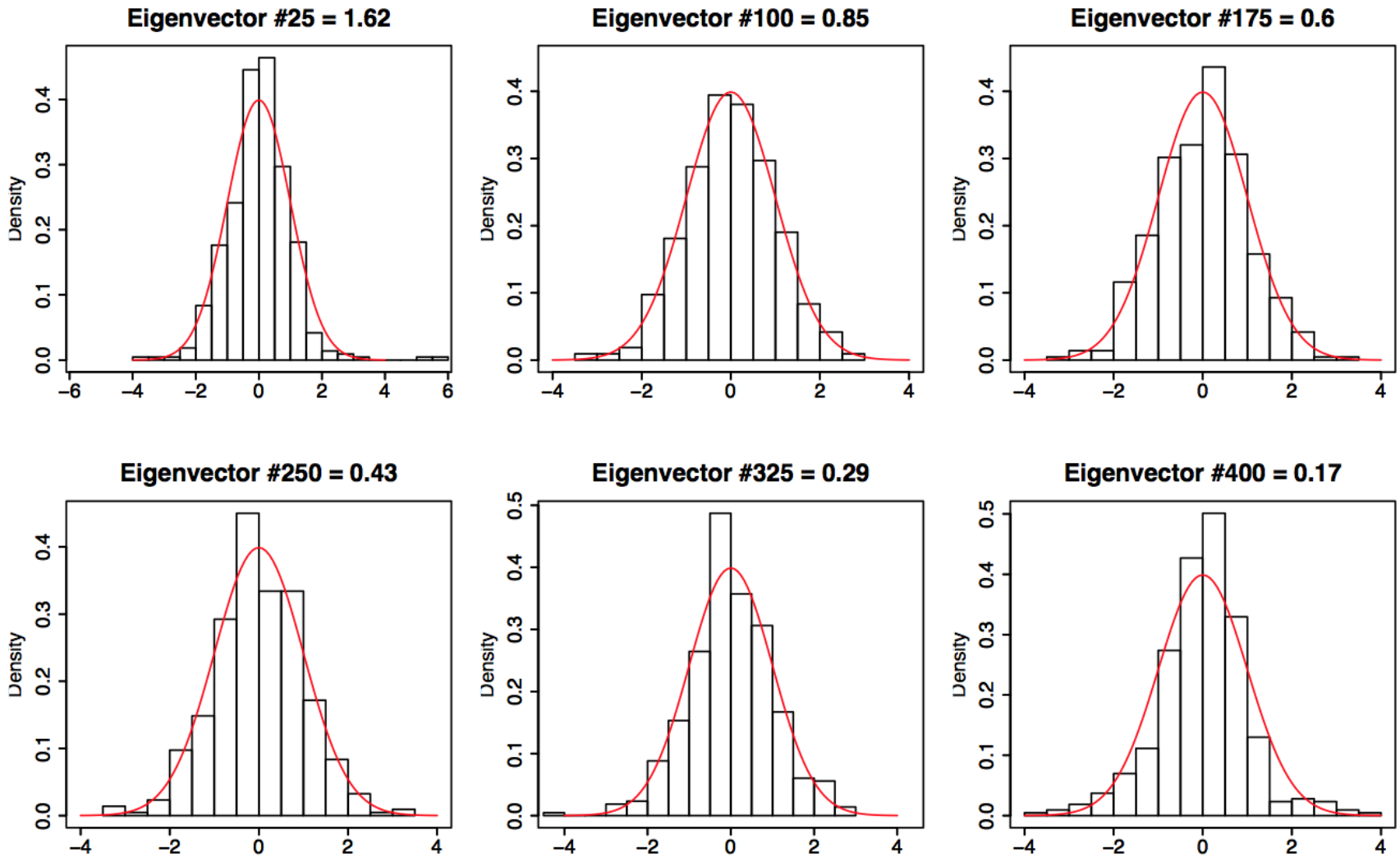
# Informative eigenvalues (1)

Plots for the six largest eigenvalues:



# Non-informative eigenvalues

Plots for six eigenvalues in the bulk of the distribution:



# Resulting recipe

1. Fit the Marcenko-Pastur distribution to the empirical density to determine $Q$ and $\sigma$
2. All eigenvalues above a threshold $\lambda^*$ are considered informative; otherwise eigenvalues relate to noise
3. Replace all noise-related eigenvalues $\lambda_i$ below $\lambda^*$ with a constant and renormalize so that $\sum_{i=1}^{M} \lambda_i = M$
   - Recall that each eigenvalue relates to the variance of a portfolio of stocks
   - A very small eigenvalue means that there exists a portfolio of stocks with very small out-of-sample variance – something we probably don't believe
4. Undo the diagonalization of the sample correlation matrix $C$ to obtain the denoised estimate $C'$
   - Remember to set diagonal elements of $C'$ to 1

# Markowitz model setup

In a single period setting, consider

- an economy consisting of $M$ risky assets (stocks)
- denote by $\ell$ and $\Sigma$ the vector of expectation and covariance matrix of the returns of the risky assets
- let $w = [w_1, \ldots, w_M]$ be the vector of weights of an investor's wealth that are invested in the risky assets.
- positive weight corresponds to long positions, while negative weight to short positions.
- if the wealth is fully invested in the risky assets, the weights sum up to 1

$$u^T w = 1,$$

where $u$ is the vector of all-ones $u = [1, 1, \ldots, 1, 1]$

- the investor's portfolio is represented by a vector of weights $w$

# Minimum variance portfolio (MVP)

- The weights $w_{mvp}$ for the minimum variance portfolio are determined by the solution to the constrained optimization problem

$$\begin{aligned} \min_{w} \quad & w^T \Sigma w \\ \text{s.t.} \quad & w^T u = 1 \end{aligned} \tag{6}$$

where $u$ denotes the all-ones vector.

- By applying the method of Lagrange multiplier, arrive at the solution

$$w_{mvp} = \frac{\Sigma^{-1} u}{u^T \Sigma^{-1} u} \tag{7}$$

- The expected return $\ell_{mvp}$ and the variance $\sigma^2_{mvp}$ of the minimum variance portfolio are given by

$$\ell_{mvp} = w_{mvp}^T \ell = \frac{u^T \Sigma^{-1} \ell}{u^T \Sigma^{-1} u} \tag{8}$$

$$\sigma^2_{mvp} = w_{mvp}^T \Sigma w_{mvp} = \frac{1}{u^T \Sigma^{-1} u} \tag{9}$$

- Note that $u^T \Sigma^{-1} u$ contains the sum of all the entries of $\Sigma^{-1}$

# Motivation

Compute and compare characteristics/performance of the minimum variance portfolios (MVP) corresponding to the

- sample covariance matrix
- filtered covariance matrix (keeping only the top 25 factors)
- Barra covariance matrix (3rd party data provider risk model)

# In-sample statistics comparison

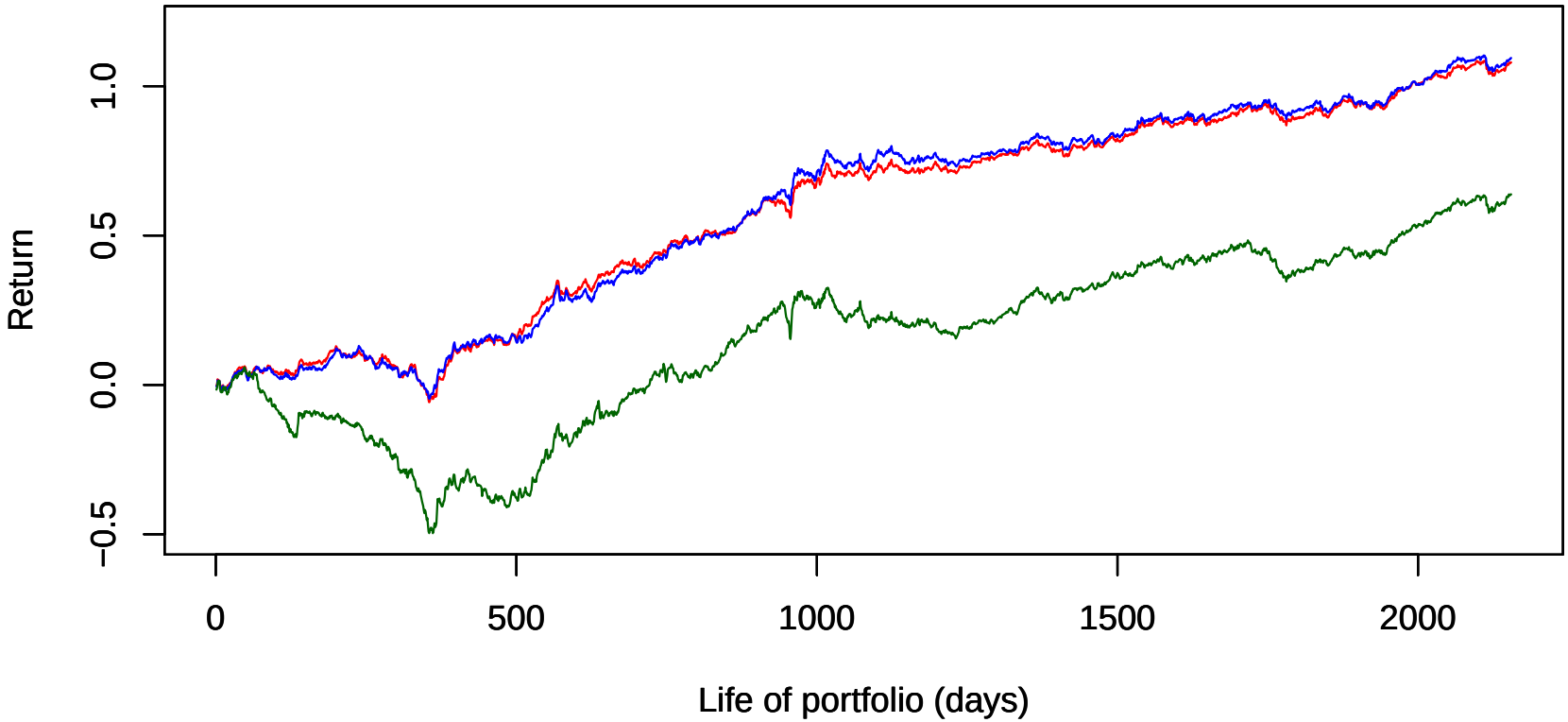


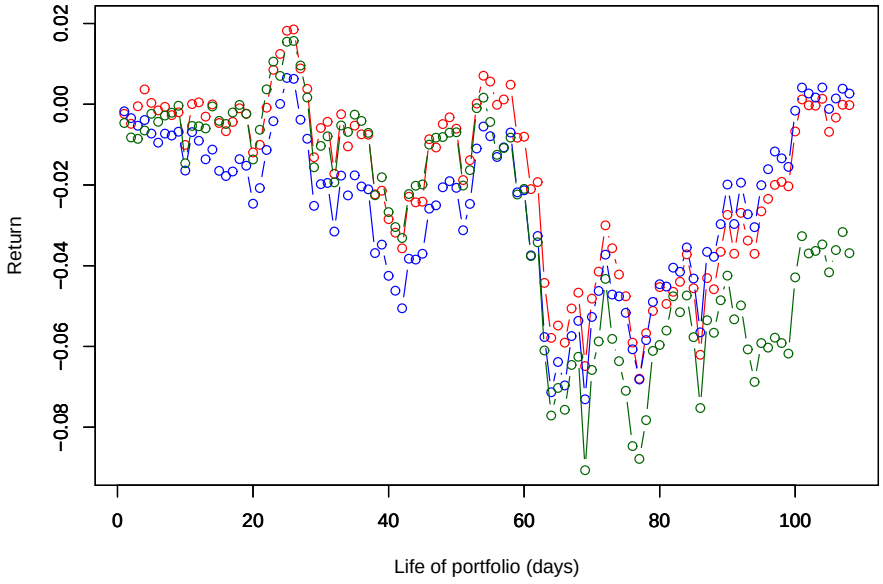

Figure: Sample in red; filtered in blue, and Barra in green.

| | Volatility | Max Drawdown |
|---|---|---|
| Sample | 0.523% | 18.8% |
| Filtered | 0.542% | 17.7% |
| Barra | 0.725% | 55.5% |

As expected, the sample portfolio has the lowest in-sample volatility.

# Out-of-sample comparison

Minimum variance portfolio returns from 04/26/2007 to 09/28/2007.




| | Volatility | Max Drawdown |
|---|---|---|
| Sample | 0.811% | 8.65% |
| Filtered | 0.808% | 7.96% |
| Barra | 0.924% | 10.63% |

Figure: Sample in red; filtered in blue, and Barra in green portfolio performances.

- The MVP computed from the RMT filtered covariance matrix wins according to both measures.
- The sample covariance matrix performs pretty well (probably because here $Q = 5$). In practice, we are likely to be dealing with more stocks (M greater) and fewer observations (T smaller).

However ...

PHYSICAL REVIEW X **4,** 031032 (2014)

# Feynman-Hellmann Theorem and Signal Identification from Sample Covariance Matrices

Lucy J. Colwell,[1] Yu Qin,[1] Miriam Huntley,[1] Alexander Manta,[2] and Michael P. Brenner[1]

[1]*School of Engineering and Applied Sciences and Kavli Institute for Bionano Science and Technology, Harvard University, Cambridge, Massachusetts 02138, USA*

[2]*Roche Diagnostics GmbH, Penzberg 82377, Germany*

(Received 17 September 2013; revised manuscript received 6 July 2014; published 27 August 2014)

A common method for extracting true correlations from large data sets is to look for variables with unusually large coefficients on those principal components with the biggest eigenvalues. Here, we show that even if the top principal components have no unusually large coefficients, large coefficients on lower principal components can still correspond to a valid signal. This contradicts the typical mathematical justification for principal component analysis, which requires that eigenvalue distributions from relevant random matrix ensembles have compact support, so that any eigenvalue above the upper threshold corresponds to signal. The new possibility arises via a mechanism based on a variant of the Feynman-Hellmann theorem, and leads to significant correlations between a signal and principal components when the underlying noise is not both independent and uncorrelated, so the eigenvalue spacing of the noise distribution can be sufficiently large. This mechanism justifies a new way of using principal component analysis and rationalizes recent empirical findings that lower principal components can have information about the signal, even if the largest ones do not.

DOI: 10.1103/PhysRevX.4.031032 Subject Areas: Statistical Physics

https://journals.aps.org/prx/abstract/10.1103/PhysRevX.4.031032

# Covariance Matrix Estimation

1. Spectrum estimation involves shrinking the sample eigenvalues while retaining the sample eigenvectors from the original matrix.
2. Structured-Based Estimation
   - imposing restrictions on the covariance matrix is a straightforward way to reduce the degrees of freedom of a covariance
   - results in a **"structured matrix"**; popular in finance
   - eg: all stock returns have the same variance and all pairs of stocks have the same covariance; factor-based models
3. Regularized-Based Estimation
   - shrinking the sample covariance matrix $S$ towards a

$$\widehat{\Sigma}(S) = \alpha T + (1 - \alpha)S,$$

     where $T$ is positive definite "target" matrix and $\alpha \in (0, 1)$ is the shrinkage intensity. Intuitively, the closer the target matrix is to the population covariance, the more intense the shrinkage should be.
   - the shrinkage intensity $\alpha$ can be determined by minimizing some desired criterion (generally estimated by cross-validation).
   - naturally leads to the classic *bias-variance trade-off* (from ML)
     - $\alpha = 1$ low-variance but high-bias estimator
     - $\alpha = 0$ low-bias but high-variance estimator

# Target $T$

Several choices are available for the shrinkage target $T$

- Perhaps the most popular choice is the ridge regularizer

$$T = \sigma^2 I \tag{10}$$

  where $\sigma^2$ is a scalar constant corresponding to the overall variance in the data
- an appropriate value is $\sigma = \text{trace}(S)/p$ so that way $S$ and $T$ have the same overall variance.

- another choice (Hoffbeck and Landgrebe) considers the diagonal components of the sample covariance

$$T = \text{diag}(S) \tag{11}$$

# Empirical analysis exercise & interesting directions to consider

- how many meaningful eigenvalues/eigenvectors (k) have there been in the US equity market over the last 10-15 years?
- sliding window based on previous m=6 months of data
- perform analysis every $s = 1$ month
- infer $k$ using RMT
- end result: plot time versus $k$, for different values of $m$
- explore the interplay with mean reversion strategies
- redo the analysis in daily traded volume space
- consider other equity markets, or other asset classes (options, bonds, crypto)
- see recent application of RMTX in options data
  - *Modeling Volatility Risk in Equity Options: a Cross-sectional approach* (talk by Marco Avellaneda) `https://math.nyu.edu/faculty/avellane/ICIBI_Volatility_2014.V2.pdf`
  - *Modeling Systemic Risk in The Options Market*, PhD thesis of Doris Dobi, Department of Mathematics, New York University 2014 `https://math.nyu.edu/faculty/avellane/DorisDobiThesis.pdf`